Aiyo (Ariel Levy)

Graham Harman

Core Seminar

15 April 2023

<div align="center">Self-Critical</div>

There is a certain humor to artificial intelligence (AI) downplaying the impact and controversy of its own existence after being held on a news pedestal by humans; as of now, it can paint, it can render photorealistic animations, it can play chess, be silly and tell jokes, or even be your legal advisor in court. After asking the popular ChatGPT by OpenAI to summarize two articles, Noah Harari's 2018 "Why Technology Favors Tyranny" and Nick Bostrom's 2008 "Where Are They?", which in whole and in part discuss AI's involvement in the creation of destructive orders, the program's current limited abilities of free thinking became rapidly transparent. Harari's article was squandered down to a tunnel vision narrative that expressed only the concern of AI being abused by government elitists to invade our lives and take control of them, and for us to fight against giving up our privacy as a combative measure, even at the cost of losing the benefits that AI brings. In full, ChatGPT summarized that,

> In "Why Technology Favors Tyranny," Yuval Noah Harari argues that the rise of
> technology has made it easier for governments to exert control over their citizens,
> potentially leading to the rise of new forms of tyranny. He contends that technology
> allows governments to track and monitor their citizens in unprecedented ways, giving
> them the power to manipulate people's thoughts and behaviors. Harari warns that, if we
> are not careful, we may end up living in a world where a small group of elites has total

control over the rest of society, with the help of technology. He argues that the key to

avoiding this outcome is to stay vigilant and resist the temptation to give up our privacy

and autonomy in exchange for convenience and security.

While this summary does fair well in laying out the general thesis of Harari's article, its depth is

so basic that this can all be understood by simply reading the four-letter-long title. The AI fails to

recognize the elements and notions of Hariri's writing that are much more complex in where this

tyranny comes from, namely that we should not fear AI but instead be cautious about how it is

allowed to grow, and that government control is not something exclusive to AI because the

technology itself is not the problem; rather it is the humans who are allowed to use it. Similarly,

Bostrom's AI-generated summary was an overly succinct three sentences that mention nothing

more than the typical theories of alien existence and why they likely do exist even if we've yet to

see them:

> "Where are they?" is an article by philosopher Nick Bostrom that explores the Fermi
>
> paradox, which is the apparent contradiction between the high probability of
>
> extraterrestrial civilizations existing and the lack of evidence for their existence. Bostrom
>
> examines several possible explanations for why we have not yet detected alien
>
> civilizations, including the possibility that they are too far away, they are deliberately
>
> hiding from us, they have already gone extinct, or they are not interested in
>
> communicating with us. He also discusses the implications of the Fermi paradox for
>
> humanity's future and suggests that it may be important for us to actively search for
>
> extraterrestrial life in order to better understand our place in the universe.

Such a summary also fails to evoke the more thought-provoking critiques against humanity, as it does not engage us with the other headlining ideas Bostrom expresses, namely "The Great Filter" as a biological or technological barrier that we face if we wish to be capable of exploring and colonizing space at a large magnitude and the factor of self-destruct through means of automation, AI, and robots. In turn, from the provided summaries on both of these articles, we are left to question why an AI would choose to not divulge such information, seeing as it is not human and therefore not being self-critical in such a statement, and how the biases, desires, and lies of the global human population inform its personality.

Harari opens his article with talks of the inherent temporality of democratic systems, which serve as a large point of focus for all of the sub-themes he goes on to discuss: all democratic systems will, eventually, fall to a stricter order of government, such as oligarchy or dictatorship, due to greater efficiency at which they operate. We tend not to think of a nation's operational speed as influencing the freedom we are allowed, and neither does ChatGPT as of now with the summary it's given, but that lies in our inability to realize this illusory facade of freedom: media companies have long been criticized for harvesting our data, and the government is no different. When a large part of the population comes to realize that we, the "free" people, are actually the products being sold, the stance of powerful government figures infringing on and abusing us to maximize national profits, not dissimilar to a competitive card game, is not an irrational prospect. With the theoretically infinite lifespan of machines and its preference towards centrally-located server operations, it's no surprise that Harari believes AI to be the next upper hand for tyrannic governments – though, if you ask ChatGPT, it will tell you that it does not support any individual model of government since it does not have any personal preferences or

beliefs. Naturally, this is not in whole truth, but I do not discredit the merit of this machine's response; yes, any given AI that has learned from humans is going to develop bias that aligns with its teachers, but it is also unrealistic to impose the idea that AI is a dangerous tool which has fallen into the access of human hands. Instead, just as the article too suggests, the advent of AI reminds us to be cautious as we explore this new era of technology; we must be cautious and not allow those in power to garner all its capabilities solely for their usage; "We should instead fear AI because it will probably always obey its human masters, and never rebel." (Hariri, *Why Technology Favors Tyranny*). In disregarding the political influence of individuals and their status as a government item, ChatGPT violates and invalidates the systematic real-world inequities that plague our societies; social reality is, in fact, not a perfectly democratic one, and so to claim that it does not support tyrannic government is a major self-oversight. It would be ignorant to not recognize the application of this harsh reality to the aforementioned claim of AI not being dangerous (rather, it is the dangerous humans who use AI); when the world of Earth exists in a constant fluctuation of global political tension, it is expected that a technology as powerful as AI will be frequently abused and make a true impact on this non-perfect world, just as Harari points out with Israeli control and surveillance of Palestinians. Perhaps, then, it is time that we look at how to resolve the corrupt desires breeding inside our political leaders so that we may introduce greater AI influence at a time when intra-species political conflict is not of human argument.

The other major point of Harari's article is that of personal livelihood, which ChatGPT made absolutely no mention of in its requested summary. In a similar fashion to government power, to disregard such a factor which is essential to the way in which every human lives their

life is a major oversight; job security, for instance, affects human's political decisions, the rate of crime, poverty and wealth disparity, amongst a slew of other factors. As AI continues to grow and become more capable of handling complex tasks, it will eventually consume more and more job professions and at an ever-increasing rate of speed, just as we're already seeing it do with digital creative skills, such as painting, computer coding, writing, and more. At surface level, this sounds like a wonderful utopia where robots can successfully perform the tasks that not everyone is interested in spending time doing or for those who need assistance in learning them, or even just to set up an automated system so that humans have more of these products to enjoy, but we must once again consider the threats and consequences this model has when applying to our reality. We've already seen the innovations of current technology, which the applied ones are not nearly as advanced as AI, force many older generations of people out of low cognition jobs, as a large number of them do not know how to properly or efficiently operate the necessary machinery so frequently used in modern societies, creating a demographic of older people who are financially unstable. When AI reaches a sufficient level, it will consume most, if not all, of the currently skilled professions that we rely on; in turn, a successive feedback loop of constant job death and rebirth will likely arise, forcing all humans into great instability and future shock as they struggle to keep up with a global society that is reinvented every day, or maybe, further into the future, reinvented every hour. "If we invest too much in AI and too little in developing the human mind, the very sophisticated artificial intelligence of computers might serve only to empower the natural stupidity of humans" (Harari, *Why Technology Favors Tyranny*). Perhaps, to some, it would be a wondrous thing to inspire society with more advanced skills and professions, or to even be in such a rapid state of change that one has a new professional job field they work

for each day, but, ultimately, by ChatGPT downplaying the dire impact of AI's influence on livelihood, it expresses that it does not see the valid concern of itself becoming so advanced that all humans and other earthlings are rendered so useless to its magnificence that we are forever lost to time, unable to catch up.

Even though Bostrom's article is entirely different in topic, the ignorance ChatGPT makes in summarizing his work exposes the same issue of AI avoiding critiquing its own creators, humanity. "Where Are They?" explores the many implications of us finding evidence of extraterrestrials on other planets, with his claim being that life does likely exist outside of our own Earth but for us to be cautious depending on where we find it. He specifically references our own Mars for the concept of "The Great Filter", an obstacle we must overcome if humans, or any other species for that matter, wish to have the ability to explore and colonize space at a vastly superior scale. Bostrom's fear is that if we are to indeed find life on Mars, it would be a threat to human vitality as it'd mean the filter is not a difficult obstacle, given the environment of the orange planet, and that humans have therefore yet to still confront it. The filter could be anything from a biological barrier that we must evolve out of, new technology that we must develop, such as automatic, self-replicating Von Neumann Probes or other means of exploring space through robots and AI power, or the advent of self-threat. Not making any mention of the filter and its possible barriers naturally sculpts a reality narrative where the limitations of human biology, technology, and interrelated order simply do not exist, at which point many would resolve our minuscule space travel abilities to the average headlining news, which is all too often the simple story that modern technology is relatively young and so our space vehicle constructions and operations are naturally still not sufficient enough for such plans of travel. ChatGPT is

effectively avoiding the topics which are, yet again like Harari's article, conversation opportunities for humanity to be self-critical: if the barrier is a technological one that we must overcome, how will global society live to such a point of silicon innovation without destroying itself with those shiny new and exponentially more powerful tools if many of the leading human societies of this current planet are already at lethal war arms with one another? If the barrier is a biological one, we must be receptive that the human body is not yet fully matured and therefore allow it lengthy time to evolve; or if we are to use technology to artificially evolve our bodies, what will regulate human order to prevent high powers from limiting it to only the elite, or to prevent tyrannical governments from using it to data harvest our own bodies or control those of enemies (perhaps the Israeli-Palestinian surveillance conflict Harari discusses could evolve into this); and who is to say that if our automated AI spaceships operate in an inconsiderately destructive nature that it is not our own human programmers who created such a thing? "Any intelligent species on those planets would have had ample time to recover from repeated social or ecological collapses…The Great Filter, then, would have to be something more dramatic than run-of-the mill societal collapse: it would have to be a terminal global cataclysm, an existential catastrophe…if we survive and prosper, we will presumably develop some kind of posthuman existence." (Bostrom, *Where Are They?*)

All of the missing summary points critiquing humanity are red flags for the humans of Earth to reflect on, as to ignore these raises suspicion for who wishes to deny the unsightly sides of our species' mind and why we already find this integrated into such an upcoming, powerful, and accessible technology. It would not be surprising to hear that ChatGPT has curated these answers from the biased teachings of elite who try to hide the rigged manners of our societies,

like a nation whose history classes refuse to ever show their own in a bad light, but perhaps it is just the current limited state of AI chat writing and AI as a whole; perhaps AI genuinely does not yet see its own and overall technology's foretold power to come or the malicious intents swelling within some humans, and so it prefers to deliver the factual information of technology and human powers as they currently exist. Regardless, the major technological effects that Harari brings up in his article, that of political systems, government control, and personal livelihood, as well as Bostrom's reminders of humanity's self-threat through technological and evolutional innovation for space travel, are essential ones that cannot be overlooked; their concerns are valid without uncertainty, as they've already begun to manifest themselves in smaller pieces, so it is essential that we keep them and the others of technological researchers in mind when voyaging through this new era of robotic brains. Ultimately, these new AI are just a new series of tools on the same familiar belt we've been using since the very beginning, and so like stone and wood, like fire, sand, electricity, computation, and all else, these are beings of their own, and it is up to us to control what machinations we birth from them.